Cognitive Mapping in Animals and Robots

Nestor Schmajuk
and
Horatiu Voicu

Department of Psychology
Duke University
Durham, NC 27707

Abstract

We describe the mapping of an unknown environment by an agent, using a Cognitive Map implemented by a neural network. The environment is depicted in terms of the connectivity between spatial places located on a Cartesian grid, or canvas. Computer simulations show that the network offers a novel description of animal behavioral paradigms such as latent learning.

Spatial Learning with a Cognitive Map

In order to map the environment, we used a Cognitive Map technique described by Schmajuk and Thieme (1992; Schmajuk, Thieme, and Blair, 1993). Schmajuk and Thieme depicted spatial learning in terms of a system composed of (a) an Action System, i.e., a goal-seeking mechanism in which different goals can be defined; and (b) a Cognitive System, i.e., a representation of the adjacency between places in the environment. The goal-seeking mechanism finds the goal by approaching (a) the goal when it is next to it, or (b) an adjacent place that best predicts, through the Cognitive Map, the optimal direction towards a still distant goal. Before the Map is built, goals are all unvisited places. Once the Map is built, goals can be defined as specific locations of special interest (a place, all places).

We applied the model to guide an autonomous agent during the survey of an unknown environment. We showed that the model can achieve an efficient and exhaustive exploration, substantially better than that achieved by random movements, even when improved by avoiding the re-visitation of places (Schmajuk and Voicu, 2000b). We have also shown that the agent is highly fault tolerant, and can achieve effective exploration when its location is not always precisely known (Voicu and Schmajuk, 2000a).

Cognitive Map. Figure 1 shows a neural network implementation of a Cognitive Map. In the Cognitive Map, $Place_i$ may become associated with the views of other places to form long-term associations $V_{i,j}$ between Place i and View j. When the system is at Place i and perceives View j, $V_{i,j}$ increases. When the system is at Place i and cannot perceive View j, $V_{i,j}$ decreases. $V_{i,j}$ is readjusted to reflect the environment configuration. The strength of these modifiable synapses are indicated by open

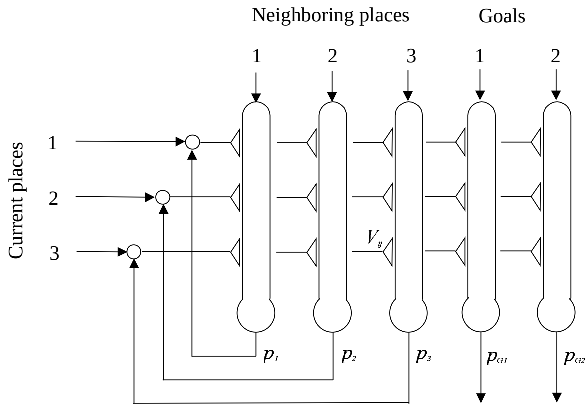

Figure 1. A neural network for Cognitive Mapping

triangles in Figure 1. Association $V_{i,j} = 1$ indicates that Places i and j are adjacent and that the agent can move from Place i to Place j. Association $V_{i,j} = 0$ can indicate either that Places i and j are not adjacent, or if they are adjacent the agent cannot move from Place i to Place j.

The Cognitive Map can combine multiple associations $V_{i,j}$ to infer spatially remote goal locations. This is achieved by recurrently reinjecting the signal representing View j as activated by Place i proportionally to $V_{i,j}$, into the representation of Place j. Place j now activates View k according to $V_{j,k}$, and the signal representing View k is re-injected into the representation of Place k. The process continues until the representation of the goal is eventually activated. The number of re-injections is proportional to the distance between places i and k. Because each re-injection attenuates the intensity of the signal representing the View, the intensity of the final activated View is inversely proportional to the number of re-injections and, therefore, to the distance between Place i and the Goal.

Building the Cognitive Map. Figure 2 shows that the empty "canvas" on which the Cognitive Map is drawn is a lattice representing the potential continuity of the space to be explored. In the figure, dashed-side squares represent places, the circles indicate their centers, and the solid lines connecting the circles represent possible movements from one place to another. The resolution of the lattice can be defined by the sensor range or footprint of the vehicle. At the beginning of the exploration,

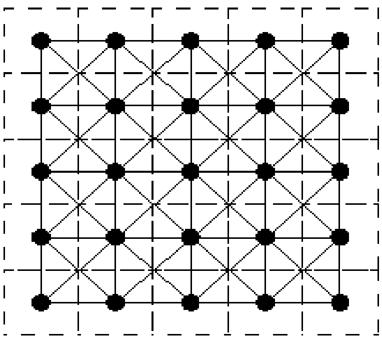

Figure 2. **The empty canvas**

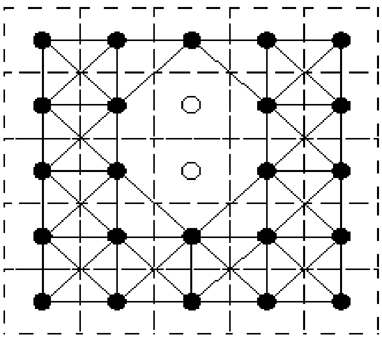

Figure 3. Barrier in the canvas

all adjacent places are assumed to be connected in the Cognitive Map, and each place is designated as "unvisited." The agent can be forced to explore a discrete region, or regions, by limiting the number of unvisited places.

Of the 8 neighboring places, the agent moves to the next place that generates the strongest prediction of "unvisited places." If more than one place generate the same strongest prediction, priority is given in the following order: North, West, East, North-West, North-East, South, South-West, and South-East. North is assumed to be the direction of the beach.

As the agent explores the environment, connections between places are modified in order to reflect the real environment. When Place j, adjacent to Place i currently occupied by the system, cannot be accessed then $V_{i,j} = V_{j,i} = 0$, and $Place_j$ remains as "unvisited." When $Place_j$, adjacent to $Place_i$ currently occupied by the system, can be accessed then $V_{i,j} = V_{j,i} = 1$ are not changed, and $Place_j$ changes its status to "visited."

In addition to the structure of the environment, described by the connectivity of the places and materialized by "barriers", "obstacles" and "goals" can also be represented on the canvas (See Figure 3).

Whereas goals represent an attractive feature, one that the system under the appropriate motivation wants to approach, an obstacle is something that can slow the progress from one place to another (represented by $V_{i,j} < 1$), and a barrier is something that completely stops progress from one place to another (represented by $V_{i,j} = 0$).

Computer Simulations

This section shows computer simulations illustrating how the model works in a paradigm known as latent learning. Blodget (1929) studied how non-rewarded trials affect performance when reward is later introduced in a multiple T-maze. A diagram of Blodget's maze is shown in Figure 4. Food-deprived rats received one trial a day, in which they were placed in the start box (Place 1) and retrieved after reaching the goal box (Place 19).

Several groups of rats were trained. One group received food reward in the goal box starting on the first trial. Two other groups were rewarded

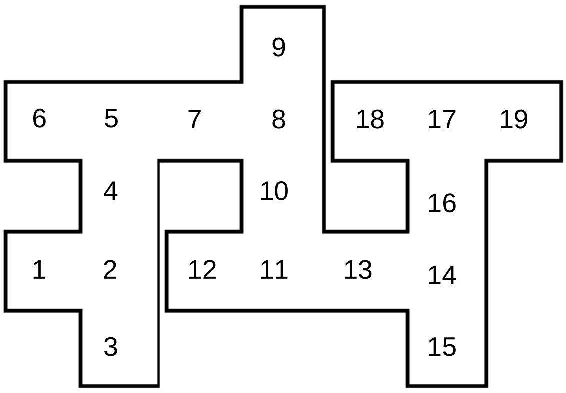

Figure 4. Diagram of a multiple-T maze employed to study latent learning.

after 3 or 7 nonrewarded trials. Blodget found that after only one rewarded trial the performance of the initially nonrewarded groups improved to nearly the same level of the group rewarded on all trials.

In our computer simulations, initially all Place-Place and Place-Goal associations are set to 1. During the exploration period, because all Place-Goal associations are 1, the agent will attempt to visit all unvisited places. Meanwhile, a Cognitive Map is constructed, which represents the maze by modifying the Place-Place associations. For instance, when the agent is in Place i and the neighboring Place j is not accessible, the Place i-Place j association is changed from 1 to 0. If the agent is located in Place i and the neighboring Place j is accessible then, the Place i-Place j association remains equal to 1. In addition, if Place j is accessible but no reward is found there, then the Place j-Goal association is will decrease gradually from 1 to 0. If reward is found in Place j, then the Place j-Goal increases gradually to the magnitude to the reward.

The simulations compared the effect of rewarding the simulated animal at the goalbox either on the first trial or on later trials. When the simulated animal is rewarded on the first trial, as the simulated animal moves through the environment, Place-Place associations are modified to reflect the maze layout. That is, if the simulated animal is in Place i and Place j if located behind a maze wall, then that Place i-Place j connection is set to 0. If instead, Place j can be accessed by the simulated animal, Place i-Place j association is unchanged. Although this process of updating the Cognitive Map takes time during the simulations, and requires time in real animals too, it is not reflected in the number of steps the simulated animal takes to reach the goal.

In addition, because the simulated animal does not find reward at any place with exception of the goalbox, Place-Goal associations are gradually decreased and goalbox-Goal associations are unchanged. Since at the beginning of experiment most of the places are still attractive, the simulated animal moves into the unvisited places and this is reflected in a larger number of steps (and a longer time) to the goal. This number gradually decrease as the attractiveness of the unrewarded places decreases. When all non rewarded places become relatively unattractive, the simulated animal moves directly to goalbox in the minimum number of steps (14).

When the simulated animal is not rewarded on the first trial, like in the previous case, as the simulated animal moves through the environment Place-Place associations are modified to reflect the maze layout. In addition, because the simulated animal does not find reward at any place, including the goalbox, all Place-Goal associations are gradually decreased. If reward is not introduced before all Place-Goal associations are reduced to 0, the simulated animal will cease exploring. If reward is introduced at the goalbox when the simulated animal is still exploring, goalbox-Goal association will increase. If by the time the reward is introduced most of the Place-Goal associations are close to 0, then the goalbox will be the only attractive place in the environment, and therefore the simulated animal will go directly to the goalbox, as illustrated by Group D in Figure 5. If, instead the Place-Goal association are not completely extinguished, then the goalbox will compete with other nonrewarded places, and therefore the simulated animal will explore the unrewarded places that are farther from the goalbox before moving towards the goalbox, as illustrated by Groups B and C in Figure 5.

Figure 5 shows the number of moves to reach the goal as a function of trials for four different groups. Group A is rewarded at the goalbox on the first trial, Group B on the eighth trial, and Group C on the tenth trial. Group D on the twelfth trial. Whereas Group A shows a gradual decrease in the number of steps, Group D shows a steep reduction once rewarded. According to Figure 5, the model shows latent learning because when reward is presented after a period of latency, animals with a pre-constructed cognitive map (Groups B, C, and D) display rapid improvement in performance to the same level of Group A. These results are in accordance with Blodget's (1929) latent learning data.

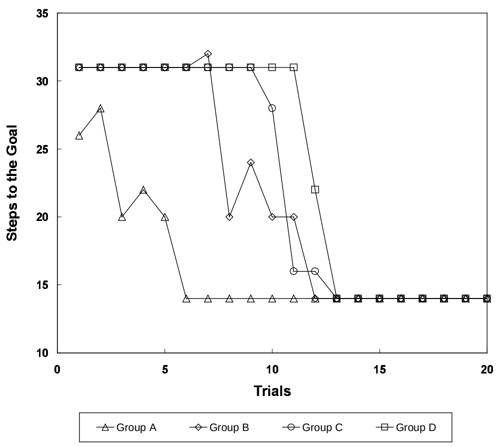

Figure 5. Latent learning. Number of simulated moves to reach the goal in a multiple-T maze for four different groups.

Discusssion

We applied a model for spatial navigation capable of guiding the search for specific goals with the assistance of a Cognitive Map. The unexplored environment is represented in the Cognitive Map as a set of directionally linked, unvisited places. During exploration, these unvisited places constitute the goals the Action System should reach. The system efficiently and completely explores the environment. Once the environment is mapped, the Action System can plan the best route to navigate between places.

When exploring an unknown environment, the system shows interesting features:

1. Achieves complete coverage (mapping) of any specified region of the environment,

2. Exploration is achieved efficiently with few revisits,

3. All inaccessible places are identified,

4. Provides reliable mapping even with unreliable dead-reckoning system and navigational signals.

The model's description of latent learning in animals substantially differs from previous descriptions. For instance, Schmajuk and Thieme (1992) assumed that in the absence of a prediction of the reward animals explore the maze by generating random (undirected) movements until (a) the Cognitive Map is built and (b) the reward is found. In contrast, the present paper assumed that maze exploration is controlled by movements directed to unvisited places according to established priorities until (a) the Cognitive Map is built, (b) the reward is found, and (c) the attractiveness of the unvisited places becomes smaller than the attractiveness of the rewarded place. One advantage of the present approach is that exploration is achieved in a more efficient and thorough manner than in the random case.

References

Blodget, Blodget, H.C. (1929). The effect of the introduction of reward upon maze performance of rats. University of California Publications in Psychology, 4, 113-134.

Schmajuk, N.A., and Thieme, A.D. Purposive behavior and cognitive mapping: An adaptive neural network. Biological Cybernetics, 67, 165-174, 1992.

Schmajuk, N.A., Thieme, A.D., and Blair, H.T. Maps, routes, and the hippocampus: A neural network approach. Hippocampus, 3, 387-400, 1993.

Voicu, H. and Schmajuk, N. (2000a) Reliable Mapping with an Unreliable Positioning System. Stanford, July 2000.

Voicu, H. and Schmajuk, N. (2000b) Three-dimensional cognitive mapping with a neural network. Robots and Autonomous Systems, in press.

Acknowledgments

This work was supported by ONR and the Dean of Trinity College at Duke University.